

# Weak Adaptive Submodularity and Group-Based Active Diagnosis with Applications to State Estimation with Persistent Sensor Faults

Sze Zheng Yong<sup>a</sup>    Lingyun Gao<sup>b</sup>    Necmiye Ozay<sup>b</sup>

**Abstract**—In this paper, we consider adaptive decision-making problems for stochastic state estimation with partial observations. First, we introduce the concept of *weak adaptive submodularity*, a generalization of adaptive submodularity, which has found great success in solving challenging adaptive state estimation problems. Then, for the problem of active diagnosis, i.e., discrete state estimation via active sensing, we show that an adaptive greedy policy has a near-optimal performance guarantee when the reward function possesses this property. We further show that the reward function for group-based active diagnosis, which arises in applications such as medical diagnosis and state estimation with persistent sensor faults, is also weakly adaptive submodular. Finally, in experiments of state estimation for an aircraft electrical system with persistent sensor faults, we observe that an adaptive greedy policy performs equally well as an exhaustive search.

## I. INTRODUCTION

The operation and control of complex cyber-physical systems often requires the choice of a sequence of decisions under partial observability. This problem of adaptive decision-making appears in various applications, both in the context of stochastic control (e.g., in robot navigation [1] and reinforcement learning [2]) and stochastic state estimation (e.g., in information gathering in robotics [3], [4], fault diagnosis in nuclear plants [5] and sensor placement [6], [7]). In particular, the research area of active diagnosis and active learning, i.e., the problem of discrete object or state identification through a minimal number of sequential actions, has many real-world applications in medical diagnosis and emergency response [8], [9], electrical systems diagnosis [10], scheduling [11], etc. Hence, progress in this direction can have a significant impact on a broad range of problems in robotics, cyber-physical systems and artificial intelligence.

**Literature Review.** Finding optimal policies for general partially observable stochastic estimation problems is often an intractable combinatorial optimization problem. Hence, researchers, e.g., [12], [13] in the context of actuator and sensor placement, have appealed to a diminishing returns property known as submodularity, which plays a similar role in combinatorial optimization as convexity in continuous optimization. Moreover, a recent paper [6] introduced the notion of adaptive submodularity for set functions that extends submodularity to the adaptive setting, i.e., when

actions are sequentially chosen based on observations, and provided near-optimal performance guarantees for adaptive greedy policies. However, there are many applications in which greedy policies perform very well even when the objective function in question is not adaptive submodular. This was observed in [14] in the non-adaptive setting, where “near-submodular” set functions are shown to also lead to near-optimal performance guarantees. The paper [15] hinted that this will hold in the adaptive case, but without providing formal justification. Thus, there is still a need for a rigorous definition of “near-submodular” set functions in the adaptive setting, and for the derivation of performance guarantees for adaptive decision-making problems with such functions.

For the problem of active diagnosis (a.k.a. adaptive stochastic maximization), where actions or queries are selected sequentially to maximize the accuracy of diagnosis with a given budget of actions, [6] showed that the reward function for this problem is adaptive monotone submodular, when outcomes are not corrupted by noise or faults. Hence, an adaptive greedy policy yields the best polynomial-time approximation algorithm [6]. However, the case when the outcomes are corrupted by noise or faults is less well studied. The papers [16], [17] considered this problem when the fault is stochastic/non-persistent, i.e., the case when repeated measurements may yield different outcomes, but the proposed algorithms have no provable performance guarantees. And to our best knowledge, the active diagnosis problem with persistent noise or faults has not been considered, although this has been investigated for the complementary problem known as active learning in [9], [18]. In these approaches, proxy set functions were found to indirectly prove the near-optimality of the original problem. However, corresponding specialized algorithms were needed and it is unclear how these apply to our active diagnosis problem. In fact, these papers [9], [18] considered a more general problem of group-based active learning, also studied in [4], of which active learning with persistent noise is a special case. The objective of group-based active learning is to learn the group to which an object or objects belong, as opposed to the objects themselves with a minimum number of actions or queries. On the other hand, the problem of group-based active diagnosis has, to the extent of our knowledge, not been investigated, despite its potentially broad applicability.

**Contributions.** To close the gap between non-adaptive and adaptive settings, we rigorously define the concept of *weak adaptive submodularity* for set functions, as a generalization of (strong) adaptive submodularity, and prove that an adaptive greedy policy yields competitive near-optimal

This work was supported by an Early Career Faculty grant from NASA’s Space Technology Research Grants Program and DARPA grant N66001-14-1-4045 and was done at the University of Michigan, Ann Arbor.

<sup>a</sup> School for Engineering in Matter, Transport and Energy, Arizona State University, Tempe, AZ, USA. (email: sze.zheng.yong@asu.edu)

<sup>b</sup> Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI, USA. (e-mail: {glingyun, necmiye}@umich.edu)

performance guarantees for the active diagnosis problem, when the objective function is adaptive monotone and weakly adaptive submodular. We observe that the adaptive submodularity factor that characterizes nearness to (strong) adaptive submodularity affects the performance bound in the same way as when approximate greedy policies are employed.

Then, we consider the *group-based active diagnosis* problem and show that the reward function corresponding to group identification is both adaptive monotone and weakly adaptive submodular; thus we also obtain competitive performance guarantees for this problem when using a simple greedy policy without the need for proxy set functions nor specialized algorithms. We believe that our approach is novel and valuable especially for its simplicity. Moreover, the adaptive submodularity factor also offers an alternative explanation for the deterioration of the performance guarantees that was observed in group-based active learning when proxy set functions are introduced (cf. [4], [9], [18]).

Furthermore, we demonstrate that the problem of *active diagnosis with persistent noise or fault* is equivalent to the group-based active diagnosis problem; hence, once again, we have provably competitive performance guarantees for the adaptive greedy policy and do not require proxy set functions and algorithms. In addition, our general formulation allows for state-dependent sensor noise or fault, and also complex faults, such as faults that give the same outcome regardless of the system states, or faults that always give the opposite outcome. Our experiments on aircraft electrical systems show that the simple adaptive greedy policy performs just as well as a brute-force policy that tries all possible actions, while only having a polynomial-time computational complexity.

## II. MOTIVATION: AIRCRAFT ELECTRICAL SYSTEM WITH PERSISTENT SENSOR FAULTS

We are motivated by the discrete state estimation problem of an aircraft electrical system via active sensing, first studied in [10] (see Figure 1 for an example of a simple circuit and the readers are referred to [10] for a detailed description of the electrical circuits). As more systems become more dependent on electric power systems, this problem is crucial to the safety of these systems. In the above work, the sensors are assumed to be *healthy* or *faultless*. But in reality, sensors can be faulty, which means that they can provide incorrect information about the unknown discrete state (i.e., operating condition or status) of the electrical components. The faults can be either stochastic or persistent, and in this paper, we are mainly interested in persistent faults. One such fault is when a faulty sensor persistently gives the opposite reading and another is when a faulty sensor always gives a constant reading (hence sometimes correctly and others incorrectly).

Despite persistent faults, our goal is similar to that in [10], i.e., to design a policy that can adaptively estimate the discrete state of the circuit by taking “actions” (i.e., opening or closing controllable contactors) and observing the sensor measurements. Note that we are only interested in estimating the unknown discrete state of the system and not the true failure mode of the sensors. Hence, if we group all possible

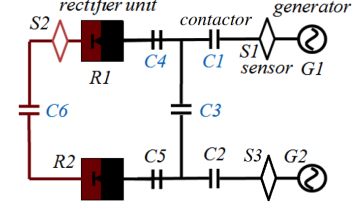


Fig. 1: A single-line diagram of a simple circuit with AC components (in black) and DC components (in red).

sensor failure modes that correspond to each state, our problem of interest is one of *group-based active diagnosis*, i.e., we aim to adaptively identify or estimate which group (i.e., state) is compatible with our sensor measurements. However, the reward function for this problem is not adaptive submodular, which motivates the need for the concept of weak adaptive submodularity, discussed next.

## III. WEAK ADAPTIVE SUBMODULARITY

In this section, we propose a generalization of the concept of adaptive submodularity and show that adaptive greedy policies achieve near-optimal performance for objective functions with such a property. These results will be useful for group-based active diagnosis, considered in the next section.

### A. Mathematical Preliminaries and Definitions

We begin by introducing notations for the adaptive decision-making problem we consider. Given a finite set of actions (or queries),  $v \in \mathcal{V}$ , our objective is to adaptively select actions to estimate a fixed but hidden state,  $\mathbf{x}_0 \in \mathcal{X}$ . We consider a Bayesian approach and model the state as a random variable,  $\mathbf{X}$ , with a given prior probability distribution  $\mathbb{P}[\mathbf{x}]$  on  $\mathcal{X}$ . When we take an action  $v$ , we observe the (sensor) measurement or outcome  $y \in \mathcal{Y}$ , which we will use to sequentially make a decision about the next action. We represent the pairs of actions  $\{v_1, \dots, v_t\}$  and observed outcomes  $\{y_1, \dots, y_t\}$  up to time  $t$  by a partial realization  $\psi_t = \{(v_i, y_i)\}_{i \in \{1, \dots, t\}}$ . Given two partial realizations  $\psi_t$  and  $\psi_{t'}$ , we call  $\psi_t$  a subrealization of  $\psi_{t'}$  if  $\psi_t \subseteq \psi_{t'}$ .

To evaluate “progress” in the decision-making process, we define an objective function  $f : 2^{\mathcal{V}} \times \mathcal{X} \rightarrow \mathbb{R}_{\geq 0}$  that maps the set of actions  $A \subseteq 2^{\mathcal{V}}$  under state  $\mathbf{x}_0$  to reward  $f(A, \mathbf{x}_0)$ . A policy  $\pi$  is a function from a set of partial realizations to actions, which specifies the action to choose next after observing  $\psi_t$ , i.e.,  $v_{t+1} = \pi(\psi_t)$ , and  $\tilde{\mathcal{V}}(\pi, \mathbf{x}_0) \subseteq 2^{\mathcal{V}}$  denotes the set of all actions under policy  $\pi$ . Randomized policies that specifies the distribution on actions are also allowed.

Next, we provide definitions specific to set functions in [6], followed by the definition of weak adaptive submodularity.

**Definition 1** (Conditional Expected Marginal Benefit). *Given an objective function  $f$ , an action  $v \in \mathcal{V}$  and a partial realization  $\psi_t$ , the conditional expected marginal benefit of  $v$  conditioned on having observed  $\psi_t$  is defined as*

$$\Delta(v|\psi_t) \triangleq \mathbb{E}[f(v_{1:t} \cup \{v\}, \mathbf{X}) - f(v_{1:t}, \mathbf{X})|\psi_t], \quad (1)$$

*with the expectation taken with respect to  $\mathbb{P}[\mathbf{x}|\psi_t]$ . Similarly, the conditional expected marginal benefit of a policy  $\pi$  is*

$$\Delta(\pi|\psi_t) \triangleq \mathbb{E}[f(v_{1:t} \cup \tilde{\mathcal{V}}(\pi, \mathbf{x}_0), \mathbf{X}) - f(v_{1:t}, \mathbf{X})|\psi_t]. \quad (2)$$

The following definition of adaptive monotonicity has the interpretation that the conditional expected marginal benefit of any action  $v$  is non-negative.

**Definition 2** (Adaptive Monotonicity). *A function  $f : 2^{\mathcal{V}} \times \mathcal{X} \rightarrow \mathbb{R}_{\geq 0}$  is adaptive monotone with respect to distribution  $\mathbb{P}[\mathbf{x}]$  if for all  $v \in \mathcal{V}$  and  $\psi_t$  with  $\mathbb{P}[\psi_t] > 0$ ,  $\Delta(v|\psi_t) \geq 0$ .*

Next, we define a new property that has the interpretation that the conditional expected marginal benefit of any fixed action (or query)  $v$  does not increase “too much” as more actions are performed and their measurements are observed.

**Definition 3** ( $\zeta$ -Weakly Adaptive Submodularity). *A function  $f : 2^{\mathcal{V}} \times \mathcal{X} \rightarrow \mathbb{R}_{\geq 0}$  is  $\zeta$ -weakly adaptive submodular with respect to distribution  $\mathbb{P}[\mathbf{x}]$  if for all  $\psi_t, \psi_{t'}$  such that  $\psi_t$  is a subrealization of  $\psi_{t'}$ , i.e.,  $\psi_t \subseteq \psi_{t'}$ , and for all  $v \in \mathcal{V} \setminus v_{1:t'}$ ,*

$$\Delta(v|\psi_{t'}) \leq \zeta \Delta(v|\psi_t),$$

for some constant  $\zeta \geq 1$  that we refer to as the adaptive submodularity factor. Additionally, let  $\zeta^* \leq \zeta$  be the smallest such  $\zeta$ , referred to as the best adaptive submodularity factor.

From the above definition, it is clear that for any  $\zeta \geq \zeta'$ , if  $f$  is  $\zeta'$ -weakly adaptive submodular, then  $f$  is also  $\zeta$ -weakly adaptive submodular. Further, (strongly) adaptive submodular functions  $f$  (with  $\zeta^* = 1$ ) are also  $\zeta$ -weakly adaptive submodular for any  $\zeta \geq 1$ ; hence, this is a generalization of (strong) adaptive submodularity. In fact, any set function is  $\zeta$ -weakly adaptive submodular with some (possibly infinite)  $\zeta$ . Note that our definition differs from approximate adaptive submodularity that was defined in [15] without justification.

### B. Adaptive Greedy Policies

Weak adaptive submodularity resembles a diminishing returns property for policies, which suggests that an adaptive greedy policy can provide nice performance guarantees. Thus, we consider an adaptive greedy policy that, at each iteration, myopically follows the greedy heuristic of maximizing its expected gain given its current observations:

$$\pi^{\text{greedy}}(\psi_t) \triangleq v_{t+1} \in \arg \max_{v \in \mathcal{V}} \Delta(v|\psi_t).$$

Such greedy strategies do not in general produce an optimal solution when compared to clairvoyant algorithms, but are oftentimes near-optimal in practice and have the advantage of having a polynomial-time computational complexity.

### C. Performance Guarantees for Active Diagnosis

We consider an important class of problems as follows:

**Problem 1** (Active Diagnosis). *Given a reward function  $f$ , find a policy  $\pi^*$  with a budget of  $k$  actions such that*

$$\begin{aligned} \pi^* &\in \arg \max_{\pi} \mathbb{E}[f(\tilde{\mathcal{V}}(\pi, \mathbf{X}), \mathbf{X})] \\ \text{subject to } |\tilde{\mathcal{V}}(\pi, \mathbf{x})| &\leq k, \forall \mathbf{x} \in \mathcal{X}, \end{aligned} \quad (3)$$

with expectation taken with respect to  $\mathbb{P}[\mathbf{x}]$  and  $\tilde{\mathcal{V}}(\pi, \mathbf{x}) \subseteq 2^{\mathcal{V}}$  is the set of all actions under policy  $\pi$  when the state is  $\mathbf{x}$ .

It turns out that for this class of problems, the weak adaptive submodularity property offers a very nice performance

guarantee when used in conjunction with adaptive greedy policies, which we will prove in Appendix A.

**Theorem 1.** *Fix any  $\zeta \geq 1$ . Let the greedy policy  $\pi_{\ell}^{\text{greedy}}$  be run for  $\ell$  iterations (so that it select  $\ell$  actions), and  $\pi_k^*$  be any policy selecting at most  $k$  actions for any realization  $\mathbf{x}$ . Then for adaptive monotone and  $\zeta$ -weakly adaptive submodular  $f$ ,*

$$f_{\text{avg}}(\pi_{\ell}^{\text{greedy}}) > (1 - e^{-\ell/\zeta k}) f_{\text{avg}}(\pi_k^*),$$

where  $f_{\text{avg}}(\pi) \triangleq \mathbb{E}[f(\tilde{\mathcal{V}}(\pi, \mathbf{X}), \mathbf{X})]$  is the expected reward of the policy  $\pi$  with respect to the distribution  $\mathbb{P}[\mathbf{x}]$ .

A corollary from the above theorem is that our greedy policy that selects  $k$  actions obtains at least  $(1 - e^{-1/\zeta})$  of the value of the optimal policy that selects  $k$  actions. Note that if  $\zeta$  is large, then the performance bound can be poor.

**Remark 1.** *The adaptive submodularity factor  $\zeta$  affects the performance guarantee in Theorem 1 in the exact way as the approximation factor of  $\alpha$ -approximate greedy policies. Just as approximate greedy policies suggest that greedy policies are robust to incorrect priors  $\mathbb{P}[\mathbf{x}]$  by a factor  $\alpha$  (cf. [6, Section 4.3] for a discussion), the  $\zeta$ -weakly adaptive submodularity property is “robust” to the deviation from (strong) adaptive submodularity by a factor  $\zeta$ .*

## IV. NEAR-OPTIMAL GROUP-BASED ACTIVE DIAGNOSIS

### A. Problem Setup

The objective of group-based active diagnosis is not to determine the unknown object but rather the group to which the object belongs via active sensing. For instance, in medical diagnosis, one is more interested to diagnose the disease that a patient has without necessarily identifying all its symptoms, while in movie recommendation systems, it may be of more interest to identify the genre that a user prefers as opposed to the exact scene preferences. As also noted in [9], the group-based diagnosis problem cannot be simply reduced to an active diagnosis problem with groups as “meta-states”, since the objects within a group do not generally yield the same observations. In our motivating example, our goal is to primarily estimate the hidden system state with no emphasis on finding the hidden persistent sensor faults. Further, under the hypotheses of different sensor failure modes for the same system state, the sensor measurements would be different.

1) *Mathematical formulation:* In this section, we further introduce notations that are specific to the group-based active diagnosis problem. We will only consider non-overlapping groups in this paper; thus, to make the connection to active diagnosis with persistent faults clear, we will refer to groups as (system) states  $x \in \mathcal{X}$  and objects in the groups as (sensor) modes  $q \in \mathcal{Q}$ . We assume that the true state  $x_0 \in \mathcal{X}$  and the true mode  $q_0 \in \mathcal{Q}$  are fixed but unknown. As before, we will consider a Bayesian approach and model the state  $X$  and mode  $Q$  as random variables with a given prior joint probability distribution  $\mathbb{P}[x, q]$  on  $\mathcal{X} \times \mathcal{Q}$ . Note that the modes need not be the same for each state.  $\mathcal{Q}$  can be taken as the union of all modes and if a particular mode  $q$  does not belong to a state  $x$ , then  $\mathbb{P}[x, q] = 0$ .

Given the true pair of state  $x_0$  and mode  $q_0$ ,  $y = \mu(v, x_0, q_0) \in \mathcal{Y}$  is the unique outcome of performing action  $v$ . Further, we define  $D(y, v, q)$  as the set of states  $x \in \mathcal{X}$  that gives the same outcome  $y$  under the action  $v$  if  $q$  were the true mode. We then define for each iteration  $t$ ,  $S_{t,q}$  as the set of all compatible states with the hypothesis that  $q$  is the true mode, i.e., the set of all states that produce the same set of outcomes  $\{\mu(v_1, x_0, q_0), \dots, \mu(v_t, x_0, q_0)\}$  under the set of actions  $\{v_1, \dots, v_t\}$ , which can also be written as

$$S_{t,q} = \cap_{i \in \{1, \dots, t\}} D(\mu(v_i, x_0, q_0), v_i, q). \quad (4)$$

Since only intersections are taken, the order of actions  $v_i$  does not matter.

2) *Problem statement:* The objective of group-based active diagnosis is to eliminate as many of the states as possible that are incompatible with observed outcomes, i.e., the uncertainty of  $X$  represented by  $\mathbb{P}[x]$ . For the group-based active diagnosis problem, this means that at any step  $t$ , we cannot eliminate any state  $x$  that is compatible with the hypothesis of one or more modes, i.e.,  $x \in \cup_{q \in \mathcal{Q}} S_{t,q}$ . Thus, we define a reward function  $f : 2^{\mathcal{V}} \times \mathcal{X} \times \mathcal{Q} \rightarrow \mathbb{R}_{\geq 0}$  as follows

$$f(v_{1:t}, \mathbf{x}) = f(v_{1:t}, x, q) = 1 + \mathbb{P}[x] - \sum_{x \in \cup_{q \in \mathcal{Q}} S_{t,q}} \mathbb{P}[x], \quad (5)$$

with  $\mathbf{x} = (x, q)$ . Note that only the final term in (5) is essential for the greedy algorithms that we propose. If  $\mathbb{P}[x]$  is uniformly distributed on  $\mathcal{X}$ , then this term is proportional to the size of  $\cup_{q \in \mathcal{Q}} S_{t,q}$ . The other terms are added such that the reward function is non-negative, and is equal to 1 when the true group  $x$  can be uniquely determined.

We also assume that we are allocated a budget on the maximum number of actions taken,  $k \ll |2^{\mathcal{V}}|$ . To achieve the goal, we want to design a policy  $\pi^*$  that finds the “best expected estimate” of the state  $x$ .

**Problem 2** (Group-Based Active Diagnosis). *Find a policy  $\pi^*$  that solves Problem 1 with  $f$  given in (5).*

### 3) Active diagnosis with persistent faults (special case):

To consider this problem as a group-based active diagnosis problem (see also [9], [18] for a similar characterization), we create groups of objects where each group corresponds to a state and the objects within the group are copies of the state for each possible mode of sensor faults. For example, we consider a system with  $|\mathcal{X}|$  states, each denoted by  $x \in \mathcal{X}$ . If we have  $m$  sensors that can be faulty and there are  $p$  ways a sensor fault can manifest itself, then we have  $(p+1)^m$  possible combinations of functioning and faulty sensor modes, denoted by  $q \in \mathcal{Q}$ , with cardinality  $|\mathcal{Q}| = (p+1)^m$ . Next, we define an object as a pair of a state and a mode, denoted as  $\mathbf{x} = (x, q) \in \mathcal{X} = \mathcal{X} \times \mathcal{Q}$ . In other words, for each group/state  $x$ , all pairs of  $\{x, q\}_{q \in \mathcal{Q}}$  form the objects corresponding to this group.

Next, to introduce the types of sensor faults we consider, we define the function  $\bar{\mu}(v, x_0)$  as the unique outcome under action  $v$  and true state  $x_0$  when all sensors are healthy and the function  $y' = \kappa_q(y)$  as the corruption of the ‘healthy’

sensor outcome  $y$  to the ‘faulty’ sensor outcome  $y'$  when the failure mode is  $q$ . Hence,  $\mu(v, x, q) = \kappa_q(\bar{\mu}(v, x))$ . Further, we define  $\bar{D}(y, v)$  with  $y = \bar{\mu}(v, x)$  as the set of states  $x \in \mathcal{X}$  that gives the same ‘healthy’ sensor measurement  $y$  under the action  $v$  when  $x$  is the actual system state. Thus,  $D(y, v, q) = \bar{D}(\kappa_q^{-1}(y), v)$ , where the inverse function  $\kappa_q^{-1}$  is such that  $y' = \kappa_q^{-1}(y)$  implies that  $y = \kappa_q(y')$ .

Note that the fault function  $\kappa_q$  is surjective, whereas  $\kappa_q^{-1}$  is an injective partial function, as can be seen with the types of sensor faults we consider in this paper. First, a Type 1 fault is when a faulty sensor persistently outputs the opposite outcome. Another type of sensor fault is when the faulty sensor always gives a constant reading, e.g., ‘0’ or ‘1’ (Type 2 fault). For a single sensor with a Type 1 fault,  $y = \kappa_q^{-1}(\kappa_q(y))$ , whereas for one with a Type 2 fault,  $y \in \kappa_q^{-1}(\kappa_q(y))$  or  $\kappa_q^{-1}(\kappa_q(y)) = \emptyset$ . To further illustrate this, consider a simple example with three sensors:  $s_1, s_2, s_3$ , where only  $s_2$  is faulty, i.e.,  $q = [1, 0, 1]$ , and the ‘faulty’ sensor measurement is  $y = [1, 0, 1]$ . If the fault is of Type 1, we have  $y' = \kappa_q^{-1}(y) = [1, 1, 1]$  and  $D(y, v, q) = \bar{D}(y', v)$ , whereas if the fault is of Type 2 (with persistent outcome ‘1’),  $\kappa_q^{-1}(y) = \{y', y''\}$ , with  $y' = [1, 0, 1]$  and  $y'' = [1, 1, 1]$ , and  $D(y, v, q) = \bar{D}(y', v) \cup \bar{D}(y'', v)$ . On the other hand, if  $s_2$  is faulty of Type 2 (with persistent outcome ‘0’) and if  $y_2$  should always be 1 for all  $q$ , then  $y' = \kappa_q^{-1}(y)$  is undefined and by definition,  $D(y, v, q) = \emptyset$ .

### B. Adaptive Greedy Policy: Group-Based Active Diagnosis

The optimal policy for Problem 2 may only be obtained by a clairvoyant algorithm. Moreover, the complexity of planning ahead for  $k$  steps scales exponentially with  $k$ . Hence, we resolve to find a scalable (polynomial-time) algorithm using an adaptive greedy strategy, i.e., to select the action at time  $t+1$  that maximizes the expected one-step reward given the available information  $\psi_t$ , denoted by  $\Delta(v|\psi_t)$  and given in Definition 1, where the expectation is taken with respect to  $\mathbb{P}[x, q|\psi_t]$ . This probability measure  $\mathbb{P}[x, q|\psi_t]$  on the set  $S_{t,q}$  for each  $q \in \mathcal{Q}$  can be obtained via Bayes’ rule:

$$\mathbb{P}[x, q|\psi_t] = \frac{\mathbb{P}[\psi_t|x, q]\mathbb{P}[x, q]}{\mathbb{P}[\psi_t]} = \begin{cases} \frac{\mathbb{P}[x, q]}{\mathbb{P}[\psi_t]} & \forall x \in S_{t,q}, \\ 0 & \text{otherwise.} \end{cases} \quad (6)$$

Note that we substituted  $\mathbb{P}[\psi_t|x, q] = 1$  if  $x \in S_{t,q}$  because the measurement process is deterministic, and  $\mathbb{P}[\psi_t|x, q] = 0$  otherwise. Moreover, since  $\sum_{q \in \mathcal{Q}} \sum_{x \in S_{t,q}} \mathbb{P}[x, q|\psi_t] = 1$ , the normalization term  $\mathbb{P}[\psi]$  can be computed as

$$\mathbb{P}[\psi_t] = \sum_{q \in \mathcal{Q}} \sum_{x \in S_{t,q}} \mathbb{P}[x, q]. \quad (7)$$

Thus, at each iteration or step  $t$ , the policy consists of choosing the next best action  $v_{t+1}$  that maximizes the gain in uncertainty reduction as defined in Definition 1, which can be rewritten for the reward function in (5) as

$$v_{t+1} \in \arg \max_{v \in \mathcal{V}} \frac{1}{\mathbb{P}[\psi_t]} \sum_{q \in \mathcal{Q}} \sum_{x \in S_{t,q}} \mathbb{P}[x, q] \left( \sum_{\tilde{x} \in \cup_{\tilde{q} \in \mathcal{Q}} S_{t,\tilde{q}}} \mathbb{P}[\tilde{x}] - \sum_{\tilde{x} \in \cup_{\tilde{q} \in \mathcal{Q}} S_{t,\tilde{q}} \cap D(\mu(v, x, q), v, \tilde{q})} \mathbb{P}[\tilde{x}] \right),$$

where  $\mathbb{P}(\psi_t) = \sum_{q \in \mathcal{Q}} \sum_{x \in S_{t,q}} \mathbb{P}[x, q]$ . Since the first term does not depend on the choice of action  $v$ , the above expression can be equivalently written as

$$v_{t+1} \in \arg \min_{v \in \mathcal{V}} \sum_{q \in \mathcal{Q}} \sum_{x \in S_{t,q}} \mathbb{P}[x, q] \sum_{\tilde{x} \in \bigcup_{\tilde{q} \in \mathcal{Q}} (S_{t,\tilde{q}} \cap D(\mu(v, x, q), v, \tilde{q}))} \mathbb{P}[\tilde{x}]. \quad (8)$$

In addition, a less computationally expensive algorithm can be obtained by further transforming (8) to

$$v_{t+1} \in \arg \min_{v \in \mathcal{V}} \sum_{y \in \mathcal{Y}} \sum_{\tilde{x} \in \bigcup_{\tilde{q} \in \mathcal{Q}} (S_{t,\tilde{q}} \cap D(y, v, \tilde{q}))} \mathbb{P}[\tilde{x}] \sum_{q \in \mathcal{Q}} \sum_{x \in S_{t,q} \cap D(y, v, q)} \mathbb{P}[x, q] \quad (9)$$

by using the fact that the sets  $\{D(y, v, q) \mid y \in \mathcal{Y}\}$  form a partition of  $\mathcal{X}$  as will be described in more detail in the proof of Lemma 2 in Appendix B.

### C. Near-Optimality Performance Guarantees

To derive performance guarantees for the group-based active diagnosis problem with the reward function  $f(v_{1:t}, x, q)$  in (5), we first show that  $f$  satisfies the following (which will be proven in Appendix B).

**Proposition 1** (Adaptive Monotonicity). *The reward function  $f(v_{1:t}, x, q)$  in (5) is adaptive monotone.*

**Proposition 2** ( $\zeta$ -Weak Adaptive Submodularity). *The reward function  $f(v_{1:t}, x, q)$  in (5) is  $\zeta$ -weakly adaptive sub-modular with*

$$1 \leq \zeta \leq \bar{\zeta} \leq \frac{|\mathcal{Q}|}{\min_{\{x \in \mathcal{X}, q \in \mathcal{Q}: \mathbb{P}[x, q] > 0\}} \mathbb{P}[x, q]}, \quad (10)$$

where  $\zeta$  and  $\bar{\zeta}$  are defined as

$$\zeta \triangleq \max_{\substack{v_{1:k} \in \mathcal{V}^k, v \in \mathcal{V}, \\ y \in \mathcal{Y}, t=1, \dots, k}} \frac{\sum_{x \in \bigcup_{q \in \mathcal{Q}} S_{t,q} \cap D(y, v, q)} \mathbb{P}[x]}{\sum_{q \in \mathcal{Q}} \sum_{x \in S_{t,q} \cap D(y, v, q)} \mathbb{P}[x, q]},$$

$$\bar{\zeta} \triangleq \max_{\substack{v_{1:k} \in \mathcal{V}^k, v \in \mathcal{V}, \\ y \in \mathcal{Y}, t=1, \dots, k}} \frac{\sum_{q \in \mathcal{Q}} \sum_{x \in S_{t,q} \cap D(y, v, q)} \mathbb{P}[x]}{\sum_{q \in \mathcal{Q}} \sum_{x \in S_{t,q} \cap D(y, v, q)} \mathbb{P}[x, q]}.$$

Both  $\zeta$  and  $\bar{\zeta}$  can be computed algorithmically, but since we are interested in a smaller  $\zeta$ , we only provide an algorithm for computing  $\zeta$  in Algorithm 1.  $\bar{\zeta}$  is of interest because it can be easily derived analytically for some special cases. Specifically, for a group-based active diagnosis problem with uniform  $\mathbb{P}[x, q]$  for each  $x \in \mathcal{X}$ , which is the case for active sensing with no informative priors on the sensor faults, we have  $\mathbb{P}[x] = |\mathcal{Q}| \mathbb{P}[x, q]$ . In this case,  $\bar{\zeta} = |\mathcal{Q}|$ . Note that these factors are only required for finding the performance guarantee, and are not needed for the adaptive greedy strategy we propose.

Since the group-based active diagnosis problem (Problem 2) is equivalent to the active diagnosis problem with an adaptive monotone and weakly adaptive submodular reward function, by Propositions 1 and 2, as well as Theorem 1, we have the following performance guarantee.

**Theorem 2.** *For any true state  $x \in \mathcal{X}$  and true mode  $q \in \mathcal{Q}$ ,*

---

### Algorithm 1 Computation of $\zeta$

---

```

1: Initialize  $\zeta = 0$  and  $S_{0,q} = \mathcal{X}$  for all  $q \in \mathcal{Q}$ 
2: for each action sequence  $v_{1:k} \in \mathcal{V}^k$  do
3:   for  $t \in \{1, \dots, k\}$  do
4:     Perform  $v_t$  and measure  $y_t = \mu(v_t, x_0, q_0)$ 
5:      $S_{t,q} = S_{t-1,q} \cap D(y_t, v_t, q)$  for all  $q \in \mathcal{Q}$ 
6:     for each action  $v \in \mathcal{V}$  do
7:       for each output  $y \in \mathcal{Y}$  do
8:          $\bar{S}_{t,q} = S_{t,q} \cap D(y, v, q)$  for all  $q \in \mathcal{Q}$ 
9:          $\bar{S}_t = \bigcup_{q \in \mathcal{Q}} \bar{S}_{t,q}$ 
10:         $\zeta_{t,v,y} = \frac{\sum_{x \in \bar{S}_t} \mathbb{P}[x]}{\sum_{q \in \mathcal{Q}} \sum_{x \in \bar{S}_{t,q}} \mathbb{P}[x, q]}$ 
11:         $\zeta \leftarrow \max\{\zeta, \zeta_{t,v,y}\}$ 
12: return  $\zeta$ 
```

---

the adaptive greedy policy  $\pi_k^{\text{greedy}}$  guarantees that

$$f_{avg}(\pi_k^{\text{greedy}}) > (1 - e^{-1/\zeta}) f_{avg}(\pi_k^*)$$

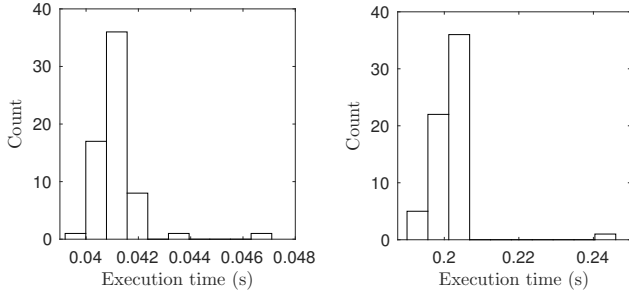
for the group-based active diagnosis problem with the reward function in (5), where  $f_{avg}(\pi_k^*)$  is what can be achieved in  $k$  steps by any other policy, including the optimal policy and  $\zeta$  is given in (10).

Hence, in the absence of informative priors on the sensor faults (i.e., when  $\bar{\zeta} = |\mathcal{Q}|$ ), our adaptive greedy policy that selects  $k$  actions obtains at least  $(1 - e^{-1/|\mathcal{Q}|})$  of the value of the optimal strategy that selects  $k$  actions. To our best knowledge, this result is novel. Group identification for the complementary problem of adaptive stochastic minimum cost cover, a.k.a. Bayesian active learning has been considered in [4], [9], [18], but not the group-based active diagnosis problem we consider in this paper. Interestingly, the performance bounds given in [9], [18] have an additional factor of 2, which is equal to  $\zeta = |\mathcal{Q}| = 2$  since they had exactly 2 modes for each persistent sensor fault, while [4] also has a factor that resembles  $\zeta$ .

### V. EXAMPLES

We tested the adaptive greedy policy on two electric power system circuits—small and large circuits. For these experiments, the state  $x$  is uniformly distributed on  $\mathcal{X}$ , and for any sensor measurement that is prone to faults,  $q_i$  for all  $i = 1, \dots, |\mathcal{Q}|$ , the conditional probabilities that faults of Type 1 and Type 2 (with outcome ‘stuck’ at ‘1’) persistently occur are  $\mathbb{P}[q_i^1 | x] = 0.8$  and  $\mathbb{P}[q_i^2 | x] = 0.6$ , respectively. The conditional probability for each state and mode pair is then  $\mathbb{P}[q | x] = \prod_{i=1}^{|\mathcal{Q}|} \mathbb{P}[q_i | x]$ . For simplicity, we only allow two states for the system components, i.e., they are either *healthy* or *faulty*. Correspondingly, the sensor readings can only take two values, i.e., *proper voltage* or *improper voltage*. We do not assume an initial action  $v_0$ , so we initialize our experiments with  $S_{0,q} = \mathcal{X}$  for each  $q$ .

For many practical cases, it is not possible to reduce the size of the compatible states to one because the number and positioning of sensors are often limited, and more so when the sensors can be faulty. Thus, we compare the performance of our adaptive greedy policy with the results of a brute force policy that exhaustively takes every action  $v \in \mathcal{V}$  (thus, violating the given budget on the number of actions). The brute



(a)  $S_2$  is fault-prone of Type 1 and (b)  $S_1, S_2$  are fault-prone of Type 1 and  $S_3$  of Type 2.

Fig. 2: Histogram of execution times (averaged over 6 actions) of the adaptive greedy policy for all true states  $x_0$ .

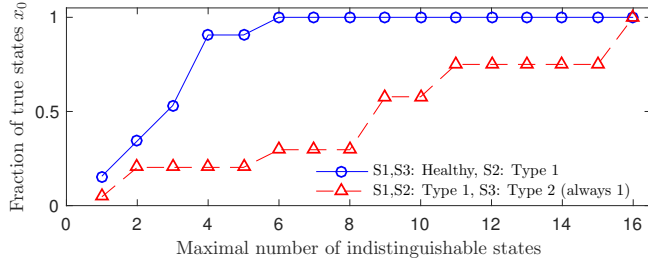


Fig. 3: Fraction of true states  $x_0$  with its corresponding maximal number indistinguishable states after 6 actions.

force policy represents the best achievable outcome, which we use as a benchmark. Both policies are run in Python on an Intel® Core™ i5-4300U 64-bit CPU@1.90GHz 2.50GHz with 8.00GB RAM, and the experiments are performed for all possible system states  $x_0 \in \mathcal{X}$  and sensor modes  $q_0 \in \mathcal{Q}$ .

#### A. Tests on a small-size circuit

We first tested the proposed adaptive greedy policy on a small circuit with 13 components, as shown in Figure 1. By using four controllable contactors  $\{C1, C3, C4, C6\}$  (thus,  $|\mathcal{V}| = 2^4 = 16$  actions) and observing the measurements of sensors  $\{S1, S2, S3\}$ , which may be faulty, the objective is to estimate the unknown states of the components  $\{G1, G2, R1, R2, C2, C5\}$ .

*Simulation result:* The adaptive greedy policy that takes 6 actions performs equally well as the brute force policy that takes all 16 actions, i.e., the values of the objective function  $f$  are equal for all possible states  $x_0$  and failure modes  $q_0$ .

1) *Average execution time:* We recorded the average execution times to find the next best action, and the results are shown in Figure 2 for two representative examples for different true sensor modes  $q$ . It takes about 25 milliseconds to compute the next best action when all the sensors are healthy, and we observed that the average execution time increases linearly in the number of modes  $|\mathcal{Q}| = 2^m$ , i.e., exponentially with the number of fault-prone sensors  $m$ .

2) *Maximal number of indistinguishable states:* Figure 3 compares the maximal numbers of indistinguishable states (corresponding to the minimal values of the reward function  $f$ ) after 6 actions for all  $x_0$  and two illustrative examples of  $q_0$ . The ordinate shows the fraction of system states  $x_0$  (out of  $2^6 = 64$ ) that obtains the number of indistinguishable states,  $|\bigcup_{q \in \mathcal{Q}} S_{\delta, q}|$ , that is less than or equal to the maximal

number (on the abscissa). Since the adaptive greedy strategy performs equally well as the brute force method, we omit the plots for the brute force policy that lie exactly on top of the plots for the adaptive greedy policy. Moreover, comparing Figures 2(a) and 2(b), we observe that the maximal number of indistinguishable states is larger when more sensors are faulty, as expected. Furthermore, for most system states  $x_0$ , the best reward is already obtained by our adaptive greedy policy within 2 actions.

#### B. Tests on a larger circuit

We also tested our greedy policy on a larger circuit with more sensors and components (cf. [10, Section V-B] for details), which is more representative of aircraft power distribution systems. When all sensors are healthy, the average execution time for obtaining the next best action is about 0.22 seconds. As before, the average execution time increases linearly with the number of sensor modes  $|\mathcal{Q}|$ . Similarly, the performance of our adaptive greedy policy with 6 actions is as good as the brute force policy with 32 actions.

## VI. CONCLUSIONS AND FUTURE DIRECTIONS

In this paper, we considered stochastic state estimation problems with partial observations via active sensing. First, we introduced a property for set functions, called *weak adaptive submodularity*, that generalizes the concept of adaptive submodularity. Then, for the active diagnosis problem, we showed that adaptive greedy policies are near-optimal when the reward function is adaptive monotone and weakly adaptive submodular. Next, we considered the group-based active diagnosis problem, for which a special case is the active diagnosis problem with persistent sensor noise or faults, and proved that the group-based reward function is weakly adaptive submodular; hence the group-based active diagnosis problem can be solved using a simple adaptive greedy policy with guaranteed competitive performance when compared to the optimal adaptive policy. Our state estimation experiments with aircraft electrical systems plagued by persistent sensor faults demonstrated that the adaptive greedy policy performs just as well as a brute-force policy.

Future work will consider group-based active learning for decision-making using weak adaptive submodularity with the hope of removing the need for proxy set functions and specialized algorithms, while preserving the provable competitiveness of adaptive greedy policies. We will also consider the general active diagnosis problem where the sensor noise and faults can be persistent and/or non-persistent.

## REFERENCES

- [1] L.P. Kaelbling, M.L. Littman, and A.R. Cassandra. Planning and acting in partially observable stochastic domains. *Artificial intelligence*, 101(1):99–134, 1998.
- [2] R.S. Sutton and A.G. Barto. *Reinforcement learning: An introduction*, volume 1. MIT Press Cambridge, 1998.
- [3] A. Singh, A. Krause, C. Guestrin, W.J. Kaiser, and M.A. Batalin. Efficient planning of informative paths for multiple robots. In *IJCAI*, volume 7, pages 2204–2211, 2007.
- [4] S. Javdani, Y. Chen, A. Karbasi, A. Krause, D. Bagnell, and S.S. Srinivasa. Near optimal bayesian active learning for decision making. In *AISTATS*, pages 430–438, 2014.



- [5] N.I. Santoso, C. Darken, G. Povh, and J. Erdmann. Nuclear plant fault diagnosis using probabilistic reasoning. In *Power Engineering Society Summer Meeting, 1999. IEEE*, volume 2, pages 714–719. IEEE, 1999.
- [6] D. Golovin and A. Krause. Adaptive submodularity: Theory and applications in active learning and stochastic optimization. *Journal of Artificial Intelligence Research*, 42:427–486, 2011.
- [7] R. Debouk, S. Lafortune, and D. Teneketzis. On an optimization problem in sensor selection. *Discrete Event Dynamic Systems*, 12(4):417–445, 2002.
- [8] T.S. Jaakkola and M.I. Jordan. Variational methods and the qmrd database. *NATO ASI Series F Computer and Systems Sciences*, 168:185–214, 1998.
- [9] G. Bellala, S.K. Bhavnani, and C. Scott. Group-based active query selection for rapid diagnosis in time-critical situations. *IEEE Transactions on Information Theory*, 58(1):459–478, 2012.
- [10] Q. Maillat, H. Xu, N. Ozay, and R.M. Murray. Dynamic state estimation in distributed aircraft electric control systems via adaptive submodularity. In *52nd IEEE Conference on Decision and Control*, pages 5497–5503. IEEE, 2013.
- [11] S.R. Kosaraju, T.M. Przytycka, and R. Borstrom. On an optimal split tree problem. In *Workshop on Algorithms and Data Structures*, pages 157–168. Springer, 1999.
- [12] T.H. Summers and J. Lygeros. Optimal sensor and actuator placement in complex dynamical networks. *IFAC Proceedings Volumes*, 47(3):3784–3789, 2014.
- [13] V. Tzoumas, A. Jadbabaie, and G.J. Pappas. Sensor placement for optimal Kalman filtering: Fundamental limits, submodularity, and algorithms. In *2016 American Control Conference*, pages 191–196, July 2016.
- [14] A. Das and D. Kempe. Submodular meets spectral: Greedy algorithms for subset selection, sparse approximation and dictionary selection. *arXiv preprint arXiv:1102.3975*, 2011.
- [15] M.J. Kusner. Approximately adaptive submodular maximization. In *NIPS workshop on Discrete Optimization and Machine Learning*, 2014.
- [16] G. Bellala, J. Stanley, S.K. Bhavnani, and C. Scott. A rank-based approach to active diagnosis. *IEEE transactions on pattern analysis and machine intelligence*, 35(9):2078–2090, 2013.
- [17] A.X. Zheng, I. Rish, and A. Beygelzimer. Efficient test selection in active diagnosis via entropy approximation. *arXiv preprint arXiv:1207.1418*, 2012.
- [18] D. Golovin, A. Krause, and D. Ray. Near-optimal bayesian active learning with noisy observations. In *Advances in Neural Information Processing Systems*, pages 766–774, 2010.

## APPENDIX

### A. Proof for Weak Adaptive Submodularity (Theorem 1)

The following lemma will be used to prove Theorem 1.

**Lemma 1.** Suppose we have observed partial realizations  $\psi_t$  after selecting  $v_{1:t}$  and let  $\pi^*$  be any policy such that  $|\mathcal{V}(\pi^*, \mathbf{x}_0)| \leq k$  for all  $\mathbf{x}_0$ . Then for adaptive monotone and  $\zeta$ -weakly adaptive submodular  $f: 2^{\mathcal{V}} \times \mathcal{X} \rightarrow \mathbb{R}_{\geq 0}$ , we have

$$\Delta(\pi^*|\psi_t) \leq \zeta \max_{A \subseteq \mathcal{V}, |A| \leq k} \sum_{v \in A} \Delta(v|\psi_t).$$

*Proof.* This result can be obtained by a slight adaptation of the proof for Lemma A.9 in [6], where we replaced  $\Delta(v|\psi_{t'}) \leq \Delta(v|\psi_t)$  with  $\Delta(v|\psi_{t'}) \leq \zeta \Delta(v|\psi_t)$  when  $\psi_{t'}$  is a subrealization of  $\psi_t$ , and with uniform action cost. ■

*Proof of Theorem 1.* The proof goes along the lines of the analysis in Theorem A.10 in [6] but for the generalized case when  $\zeta \geq 1$  and with uniform action cost. First, we consider a corollary of Lemma 1 for any  $i$ , which allows us to bound

$$\mathbb{E}[\Delta(\pi_k^*|\psi_i)] \leq \zeta k \max_{v \in \mathcal{V}} (\Delta(v|\psi_i)),$$

where the expectation is taken over the internal randomness of  $\pi_k^*$ , if there is any, and we know that the budget of actions

for all  $\mathbf{x}_0$  is such that  $\mathbb{E}[|\mathcal{V}(\pi^*, \mathbf{x}_0)|] \leq k$ . Then,

$$\begin{aligned} f_{avg}(\pi_{i+1}) - f_{avg}(\pi_i) &= \max_{v \in \mathcal{V}} (\Delta(v|\psi_i)) \\ &\geq \mathbb{E}[\Delta(\pi_k^*|\psi_i)] / \zeta k \\ &= (f_{avg}(\pi_i @ \pi_k^*) - f_{avg}(\pi_i)) / \zeta k, \end{aligned}$$

where  $\pi_i @ \pi_k^*$  is a policy obtained by running  $\pi_i$  to completion, and then running policy  $\pi_k^*$  as if from a fresh start and ignoring the gathered information  $\psi_i$ . The first equality holds because  $\pi_i$  (superscript omitted for brevity) is a greedy policy, while the final equality holds by taking an appropriate convex combination over the random partial realizations  $\Psi_i$ .

Moreover, due to the adaptive monotonicity of  $f$ , by [6, Lemma A.8], we have  $f_{avg}(\pi_k^*) \leq f_{avg}(\pi_i @ \pi_k^*)$ . Putting them together, we have for all  $i = 0, \dots, \ell$ ,

$$f_{avg}(\pi_k^*) - f_{avg}(\pi_i) \leq \zeta k (f_{avg}(\pi_{i+1}) - f_{avg}(\pi_i)).$$

Then, defining  $\Delta_i \triangleq f_{avg}(\pi_k^*) - f_{avg}(\pi_i)$ , we obtain  $\Delta_{i+1} \leq (1 - \frac{1}{\zeta k}) \Delta_i$  and thus,  $\Delta_\ell \leq (1 - \frac{1}{\zeta k})^\ell \Delta_0 < e^{-\ell/\zeta k} \Delta_0$ , where we used the fact that  $1 - x < e^{-x}$  for all  $x$  to obtain the last inequality. Finally, by the definition of  $\Delta_i$ ,

$$\begin{aligned} f_{avg}(\pi_k^*) - f_{avg}(\pi_\ell) &< e^{-\ell/\zeta k} (f_{avg}(\pi_k^*) - f_{avg}(\pi_0)) \\ &\leq e^{-\ell/\zeta k} f_{avg}(\pi_k^*), \end{aligned}$$

since  $f_{avg}(\pi_0) \geq 0$  by assumption. Hence, the result in the theorem follows directly from the above. ■

### B. Proofs for Near-Optimal Group-Based Active Diagnosis

1) *Conditional Expected Marginal Benefit*  $\Delta(v|\psi_t)$ : Our proofs for Propositions 1 and 2 rely heavily on the particular form that  $\Delta(v|\psi_t)$  takes on, given in the following:

**Lemma 2** (Conditional Expected Marginal Benefit). *The conditional expected marginal benefit of action  $v$  for the reward function  $f(v_{1:t}, x, q)$  in (5) can be expressed as*

$$\Delta(v|\psi_t) = \sum_{i=1}^{|\mathcal{V}|} \zeta_i \tau_i - \frac{\sum_{i=1}^{|\mathcal{V}|} \zeta_i \tau_i^2}{\sum_{i=1}^{|\mathcal{V}|} \tau_i}, \quad (11)$$

for some  $\tau_i \geq 0$  and  $\zeta_i \geq 1$  that are both explicitly dependent on the partial realization  $\psi_t$  for all  $1 \leq i \leq |\mathcal{V}|$ .

*Proof.* By definition,

$$\begin{aligned} \Delta(v|\psi_t) &= \mathbb{E}[f(v_{1:t} \cup \{v\}, X, Q) - f(v_{1:t}, X, Q) | \psi_t] \\ &= \mathbb{E}[\tilde{f}(v_{1:t} \cup \{v\}, X, Q) - \tilde{f}(v_{1:t}, X, Q) | \psi_t] \end{aligned}$$

with the expectation taken with respect to  $\mathbb{P}[x, q | \psi_t]$  and where  $\tilde{f}(v_{1:t}, x, q) \triangleq -\sum_{x \in \bigcup_{q \in \mathcal{Q}} S_{t,q}} \mathbb{P}[x]$ . The individual terms can be computed as:

$$\begin{aligned} &\mathbb{E}[\tilde{f}(v_{1:t} \cup \{v\}, X, Q) | \psi_t] \\ &= \sum_{q \in \mathcal{Q}} \sum_{x \in S_{t,q}} \mathbb{P}[x, q | \psi_t] \tilde{f}(v_{1:t} \cup \{v\}, x, q) \\ &= - \sum_{q \in \mathcal{Q}} \sum_{x \in S_{t,q}} \mathbb{P}[x, q | \psi_t] \sum_{\tilde{x} \in \bigcup_{\tilde{q} \in \mathcal{Q}} (S_{t,\tilde{q}} \cap D(\mu(v, x, q), v, \tilde{q}))} \mathbb{P}[\tilde{x}], \end{aligned} \quad (12)$$

$$\begin{aligned} \mathbb{E}[\tilde{f}(v_{1:t}, X, Q) | \psi_t] &= \sum_{q \in \mathcal{Q}} \sum_{x \in S_{t,q}} \mathbb{P}[x, q | \psi_t] \tilde{f}(v_{1:t}, x, q) \\ &= - \sum_{\tilde{x} \in \bigcup_{\tilde{q} \in \mathcal{Q}} S_{t,\tilde{q}}} \mathbb{P}[\tilde{x}], \end{aligned} \quad (13)$$

where the final equality holds because  $\tilde{f}(v_{1:t}, x, q)$  is independent of  $q$  and  $x$ .

To proceed, since the probability measure on  $(\mathcal{X}, \mathcal{Q})$  is non-uniform and can take values in some set  $\{p_1, \dots, p_N\}$ , we define subsets of  $(\mathcal{X}, \mathcal{Q})$  where  $\mathbb{P}[x, q]$  is constant with  $F_{n,q} \triangleq \{(x, q) \in \mathcal{X} \times \mathcal{Q} \mid \mathbb{P}[x, q] = p_n\}$ . Furthermore, since for every  $x \in \mathcal{X}$ , there is a corresponding  $y = \mu(v, x, q)$  for any given  $v \in \mathcal{V}$  and  $q \in \mathcal{Q}$ , the sets  $\{D(y, v, q) \mid y \in \mathcal{Y}\}$  also form a partition of  $\mathcal{X}$ . Finally, we define a partition of  $S_{t,q}$  for each  $q \in \mathcal{Q}$  using the sets  $\{S_{t,q} \cap D(y, v, q) \cap F_{n,q} \mid y \in \mathcal{Y}, n \in 1 : N\}$  and denote with a shorthand  $\alpha_{n,y,q} \triangleq S_{t,q} \cap D(y, v, q) \cap F_{n,q}$ . Thus, for all  $(x, q) \in \alpha_{n,y,q}$ ,  $\mu(v, x, q) = y$  and  $\mathbb{P}[x, q] = p_n$ .

Moreover, for all  $x \in F_{n,q}$ , the conditional probabilities in (6) can be rewritten as

$$\mathbb{P}[x, q \mid \psi_t] = \frac{p_n}{\mathbb{P}[\psi_t]} \triangleq \frac{p_n}{\sum_{y \in \mathcal{Y}} \tau_y}, \quad (14)$$

where we obtain the expression for  $\mathbb{P}[\psi_t]$  from (7) as

$$\begin{aligned} \mathbb{P}[\psi_t] &= \sum_{q \in \mathcal{Q}} \sum_{x \in S_{t,q}} \mathbb{P}[x, q] \\ &= \sum_{n \in 1:N} \sum_{y \in \mathcal{Y}} \sum_{q \in \mathcal{Q}} \sum_{x \in S_{t,q} \cap D(y, v, q) \cap F_{n,q}} p_n \\ &= \sum_{y \in \mathcal{Y}} \sum_{n \in 1:N} \sum_{q \in \mathcal{Q}} p_n |\alpha_{n,y,q}| = \sum_{y \in \mathcal{Y}} \tau_y, \end{aligned}$$

and defined  $\tau_y \triangleq \sum_{n \in 1:N} \sum_{q \in \mathcal{Q}} p_n |\alpha_{n,y,q}| \geq 0$ .

Furthermore, the following inequality holds:

$$\sum_{x \in \bigcup_{q \in \mathcal{Q}} S_{t,q}} \mathbb{P}[x] = \sum_{q \in \mathcal{Q}} \sum_{x \in \bigcup_{q \in \mathcal{Q}} S_{t,q}} \mathbb{P}[x, q] \geq \sum_{q \in \mathcal{Q}} \sum_{x \in S_{t,q}} \mathbb{P}[x, q], \quad (15)$$

since on the right hand side,  $q$  is a particular choice of  $\tilde{q} \in \mathcal{Q}$ . Using the above-defined partitions, we can equivalently write

$$\begin{aligned} \sum_{x \in \bigcup_{q \in \mathcal{Q}} S_{t,q} \cap D(y, v, \tilde{q})} \mathbb{P}[x] &= \sum_{n \in 1:N} \sum_{x \in \bigcup_{q \in \mathcal{Q}} S_{t,q} \cap D(y, v, \tilde{q}) \cap F_{n,q}} \mathbb{P}[x] \\ &= \zeta(\psi_t, y) \sum_{n \in 1:N} \sum_{q \in \mathcal{Q}} \sum_{x \in \alpha_{n,y,q}} p_n \\ &= \zeta(\psi_t, y) \tau_y, \end{aligned}$$

where we defined

$$\zeta(\psi_t, y) \triangleq \frac{\sum_{x \in \bigcup_{q \in \mathcal{Q}} S_{t,q} \cap D(y, v, \tilde{q})} \mathbb{P}[x]}{\sum_{q \in \mathcal{Q}} \sum_{x \in S_{t,q} \cap D(y, v, q)} \mathbb{P}[x, q]} \quad (16)$$

and from (15), we have  $\zeta(\psi_t, y) \geq 1$ . This simplifies (12) to

$$\begin{aligned} \mathbb{E}[\tilde{f}(v_{1:t} \cup \{v\}, X, Q) \mid \psi_t] &= - \sum_{q \in \mathcal{Q}} \sum_{x \in S_{t,q}} \mathbb{P}[x, q \mid \psi_t] \zeta(\psi_t, \mu(v, x, q)) \tau_{\mu(v, x, q)} \\ &= - \frac{1}{\mathbb{P}[\psi_t]} \sum_{y \in \mathcal{Y}} \sum_{n \in 1:N} \sum_{q \in \mathcal{Q}} \sum_{x \in \alpha_{n,y,q}} \mathbb{P}[x, q] \zeta(\psi_t, y) \tau_y \\ &= - \frac{\sum_{y \in \mathcal{Y}} \zeta(\psi_t, y) \tau_y^2}{\sum_{y \in \mathcal{Y}} \tau_y}. \end{aligned} \quad (17)$$

Similarly, (13) simplifies to

$$\begin{aligned} \mathbb{E}[\tilde{f}(v_{1:t}, X, Q) \mid \psi_t] &= - \sum_{y \in \mathcal{Y}} \sum_{\tilde{x} \in \bigcup_{q \in \mathcal{Q}} S_{t,q} \cap D(y, v, \tilde{q})} \mathbb{P}[\tilde{x}] \\ &= - \sum_{y \in \mathcal{Y}} \zeta(\psi_t, y) \tau_y. \end{aligned} \quad (18)$$

Finally, from (17) and (18), we obtain  $\Delta(v \mid \psi_t)$  of the form (11) with  $\zeta_y = \zeta(\psi_t, y) \geq 1$  and  $\tau_y \geq 0$  for all  $y \in \mathcal{Y}$ . ■

2) *Performance Guarantees:* Armed with Lemma 2, Proposition 1 can be shown to directly hold.

*Proof of Proposition 1.* From (11) in Lemma 2,

$$\begin{aligned} \Delta(v \mid \psi_t) &= \frac{(\sum_{i=1}^{|\mathcal{Y}|} \tau_i)(\sum_{i=1}^{|\mathcal{Y}|} \zeta_i \tau_i) - \sum_{i=1}^{|\mathcal{Y}|} \zeta_i \tau_i^2}{\sum_{i=1}^{|\mathcal{Y}|} \tau_i} \\ &= \frac{\sum_{i=1}^{|\mathcal{Y}|} (\zeta_i \tau_i \sum_{j \neq i} \tau_j)}{\sum_{i=1}^{|\mathcal{Y}|} \tau_i} \geq 0, \end{aligned}$$

since  $\zeta_i \geq 1$  and  $\tau_i \geq 0$  for all  $i = 1, \dots, |\mathcal{Y}|$ . Hence, adaptive monotonicity holds by Definition 2. ■

To prove Proposition 2, i.e., that the reward function  $f$  in (5) is weakly adaptive submodular, we first state the following lemma from [10] that will be useful.

**Lemma 3.** [10, Lemma 1] *The function  $b : \mathbb{R}^{|\mathcal{Y}|} \rightarrow \mathbb{R}$*

$$b(\tau_1, \tau_2, \dots, \tau_{|\mathcal{Y}|}) = \sum_{i=1}^{|\mathcal{Y}|} \tau_i - \frac{\sum_{i=1}^{|\mathcal{Y}|} \tau_i^2}{\sum_{i=1}^{|\mathcal{Y}|} \tau_i} \quad (19)$$

*is increasing on the positive orthant, i.e.,  $b(\tau_1, \tau_2, \dots, \tau_{|\mathcal{Y}|}) \geq b(s_1, s_2, \dots, s_{|\mathcal{Y}|})$  if  $\tau_i \geq s_i \geq 0$  for all  $1 \leq i \leq |\mathcal{Y}|$ .*

*Proof of Proposition 2.* First, taking the partial derivative of (11) with respect to  $\zeta_i$ , we obtain

$$\frac{\partial \Delta(v \mid \psi_t)}{\partial \zeta_i} = \tau_i - \frac{\tau_i^2}{\sum_{i=1}^{|\mathcal{Y}|} \tau_i} = \frac{\tau_i \sum_{j=1, \dots, |\mathcal{Y}|, j \neq i} \tau_j}{\sum_{i=1}^{|\mathcal{Y}|} \tau_i} \geq 0.$$

From the above, we can conclude that

$$b(\psi_t) \leq b(\psi_t) \min_i \zeta_i \leq \Delta(v \mid \psi_t) \leq b(\psi_t) \max_i \zeta_i \leq \zeta b(\psi_t), \quad (20)$$

where we used  $b(\psi_t)$  as a shorthand for  $b(\tau_1, \tau_2, \dots, \tau_{|\mathcal{Y}|})$  defined in (19) since  $\psi_t$  is related to  $\tau_i$ 's via (11), and  $\zeta$  is given in (10), which is conservatively chosen to be larger than  $\max_i \zeta_i$  across all possible realizations and sequence of actions, while by Lemma 2, we have  $\min_i \zeta_i \geq 1$ .

Moreover, the expression for  $\Delta(v \mid \psi_{t'})$  is similar for the partial realization  $\psi_{t'}$  but with different  $\tau'_i$  and  $\zeta'_i$ . For  $\psi_t$  as a subrealization of  $\psi_{t'}$ , i.e.,  $\psi_t \subseteq \psi_{t'}$ , by construction,  $S_{t',q} \subseteq S_{t,q}$  for all  $q \in \mathcal{Q}$  and by definition of  $\tau_i$  and Lemma 3, we know that  $\tau'_i \leq \tau_i$  for all  $i$  and  $b(\psi_{t'}) \leq b(\psi_t)$ .

Note that the following inequality from (20)

$$\Delta(v \mid \psi_{t'}) \leq \zeta b(\psi_{t'})$$

also holds for  $\psi_{t'}$  since we have chosen  $\zeta$  to be the maximum across all  $t$ . Thus, from the above and from (20), we find

$$\Delta(v \mid \psi_t) \geq b(\psi_t) \geq b(\psi_{t'}) \geq \frac{1}{\zeta} \Delta(v \mid \psi_{t'}),$$

from which  $\zeta$ -weakly adaptive submodularity holds by Definition 3. Further, we can find the upper bounds on  $\zeta$  since

$$\begin{aligned} \sum_{x \in \bigcup_{q \in \mathcal{Q}} S_{t,q} \cap D(y, v, q)} \mathbb{P}[x] &\leq \sum_{q \in \mathcal{Q}} \sum_{x \in S_{t,q} \cap D(y, v, q)} \mathbb{P}[x] \\ &\leq \sum_{q \in \mathcal{Q}} \sum_{x \in \mathcal{X}} \mathbb{P}[x] = |\mathcal{Q}| \end{aligned}$$

and  $\sum_{q \in \mathcal{Q}} \sum_{x \in S_{t,q} \cap D(y, v, q)} \mathbb{P}[x, q] \geq \min_{\{x \in \mathcal{X}, q \in \mathcal{Q} : \mathbb{P}[x, q] > 0\}} \mathbb{P}[x, q]$ , while  $\zeta \geq 1$  by (16). ■